**Automatic Detection for Passive Acoustic Monitoring of the African Elephant**

Jonathan Gomes-Selman1\*, Nikita Demir1\*, Andreas Paepcke1

1Stanford University, \*The two first authors made equal contributions

**Background/Questions/Methods (190 words):**

Poaching, illegal logging, infrastructure development, and other human activities present great threats to wildlife, demanding increased conservation and monitoring efforts. For species that inhabit vast territories or exist in remote areas, monitoring tasks become much more difficult. One promising solution is Passive Acoustic Monitoring (PAM), which involves placing autonomous recording devices throughout a habitat to passively record the vocalizations of species. However, extracting meaningful semantics from large scale audio datasets present unique non-trivial challenges that, to be effective, require automation. Herein we focus our attention primarily on the African forest elephant, deeply threatened by human behaviors, such as poaching and deforestation, which have led to dramatic decreases in the population -- over 60% in the last 100 years. Through the Cornell Lab of Ornithology, extensive acoustic data has been collected with the goal of better understanding population distributions, behaviors, and threats to survival. We employ state of the art deep learning and signal processing approaches to automate the process of per time-step audio segmentation for the identification of elephant calls. In addressing this problem, we also explore new methods for tackling the heavy data imbalance, a challenge prevalent in passive acoustic monitoring.

**Results/Conclusions (198 words):**

The task of automatically identifying elephant calls is extremely useful to conservationists who have hundreds of thousands of hours of audio data and need to segment out the sections with elephant calls locally. Without an automated approach conservationists have to resort to expensive and time-intensive human labelling. Initial results from our convolutional bidirectional long short-term memory (LSTM) recurrent neural network (RNN) model shows improvement on the state of the art for detecting elephant calls. Specifically, our model cuts down on the amount of data that human annotators need to inspect by 98%, while ensuring that approximately 96% of all actual elephant calls are correctly identified. To further improve our results and cut on the amount of human supervision for false positive elimination, we look primarily to tackle issues in large scale data imbalance and data modeling. We employ a digital signal preprocessing approach to eliminate background noise, simplifying the learning task for our model. Additionally, we explore an algorithmic method to deal with large scale data imbalance through the use of a dynamically weighted focal loss function, as well as a novel two stage approach for informatively undersampling negative background data to highlight challenging negative samples for learning.